UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/801,420	03/16/2004	Charu C. Aggarwal	YOR920040039US1	2046

7590 03/09/2007

Ryan, Mason & Lewis, LLP
90 Forest Avenue
Locust Valley, NY 11560

EXAMINER	
LOVEL, KIMBERLY M	
ART UNIT	**PAPER NUMBER**
2167	

SHORTENED STATUTORY PERIOD OF RESPONSE	MAIL DATE	DELIVERY MODE
3 MONTHS	03/09/2007	PAPER

Please find below and/or attached an Office communication concerning this application or proceeding.

If NO period for reply is specified above, the maximum statutory period will apply and will expire 6 MONTHS from the mailing date of this communication.

Office Action Summary	Application No.	Applicant(s)	
	10/801,420	AGGARWAL ET AL.	
	Examiner	Art Unit	
	Kimberly Lovel	2167	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☒ Responsive to communication(s) filed on *06 December 2006*.

2a)☐ This action is **FINAL**. 2b)☒ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) *1-31* is/are pending in the application.

4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) *1-31* is/are rejected.

7)☐ Claim(s) _____ is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9)☐ The specification is objected to by the Examiner.

10)☐ The drawing(s) filed on _____ is/are: a)☐ accepted or b)☐ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☐ All b)☐ Some * c)☐ None of:

1.☐ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____.

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) ☒ Notice of References Cited (PTO-892)
2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3) ☐ Information Disclosure Statement(s) (PTO/SB/08) Paper No(s)/Mail Date _____.
4) ☐ Interview Summary (PTO-413) Paper No(s)/Mail Date. _____ .
5) ☐ Notice of Informal Patent Application
6) ☐ Other: _____.

DETAILED ACTION

1. This communication is responsive to the Amendment filed 7 December 2007.

2. Claims 1-31 are pending in this application. Claims 1, 16 and 31 are independent. In the Amendment filed 7 December 2007, claims 1, 5, 15, 16, 20, 30 and 31 have been amended. This action is made Non-Final due to the introduction of a new rejection under 35 U.S.C 101.

3. The rejections of claims 1-5, 8, 9, 11-20, 23, 24 and 26-31 as being unpatentable over US Patent No. 6,947,933 to Smolsky in view of US PGPub 2004/0098617 to Sekar et al and claims 6, 7, 10, 21, 22 and 25 as being unpatentable over US Patent No. 6,947,933 to Smolsky in view of US PGPub 2004/0098617 to Sekar et al in view of US Patent No 6,625,585 to MacCuish et al have been withdrawn.

Claim Rejections - 35 USC § 101

4. 35 U.S.C. 101 reads as follows:

> Whoever invents or discovers any new and useful process, machine, manufacture, or composition of matter, or any new and useful improvement thereof, may obtain a patent therefor, subject to the conditions and requirements of this title.

5. The rejections of **claims 1-15** because the claimed invention is directed to non-statutory subject matter have been withdrawn as necessitated by amendment.

6. **Claim 31** is rejected under 35 U.S.C. 101 because the claimed invention is directed to non-statutory subject matter.

Claim 31 recites an article of manufacture for monitoring abnormalities in a data stream, comprising a machine readable medium containing one or more programs which when executed implement steps.

This claimed subject matter lacks a practical application of a judicial exception (law of nature, abstract idea, naturally occurring article/phenomenon) since it fails to produce a useful, concrete and tangible result.

Specifically, the claimed subject matter does not produce a tangible result because the claimed subject matter fails to produce a result that is limited to having real world value rather than a result that may be interpreted to be abstract in nature as, for example, a thought, a computation, or manipulated data. More specifically, the claimed subject matter provides for changing values of the node if the node exists. However, it is unclear what the tangible result is if the node does not exist and thus, fails to achieve the required status of having real world value.

It is suggested that claim 31 be amended to be consistent with claim 1, since the method of claim 1 provides a tangible result.

To allow for compact prosecution, the examiner will apply prior art to these claims as best understood, with the assumption that applicant will amend to overcome the stated 101 rejections.

Claim Rejections - 35 USC § 103

7. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

8. This application currently names joint inventors. In considering patentability of the claims under 35 U.S.C. 103(a), the examiner presumes that the subject matter of the various claims was commonly owned at the time any inventions covered therein were made absent any evidence to the contrary. Applicant is advised of the obligation under 37 CFR 1.56 to point out the inventor and invention dates of each claim that was not commonly owned at the time a later invention was made in order for the examiner to consider the applicability of 35 U.S.C. 103(c) and potential 35 U.S.C. 102(e), (f) or (g) prior art under 35 U.S.C. 103(a).

9. Claims 1-6, 9, 12, 16-21, 24 and 27 are rejected under 35 U.S.C. 103(a) as being unpatentable over US PGPub 2002/0161763 to Ye et al (hereafter Ye) in view of US PGPub 2002/0107858 to Lundahl et al (hereafter Lundahl).

Referring to claim 1, Ye discloses a method for monitoring abnormalities in a data stream (see abstract and [0030]), comprising the steps of:

receiving a plurality of objects in the data stream [stream of data] (see [0035], lines 5-8);

creating one or more clusters from the plurality of objects (see [0035], lines 10-13), wherein at least a portion of each of the one or more clusters comprises statistical data [sample variance, sample covariance and sample mean] representative of the respective cluster (see [0041]);

Ye discloses clustering objects and determining if an object is abnormal compared to a distance value (see [0157]-[0170]), however, Ye fails to explicitly disclose the further limitations of determining from the statistical data whether each of

the one or more clusters is abnormal when compared to a predefined value and reporting at least one of the one or more clusters as an abnormal cluster of objects in the data stream. Lundahl discloses performing cluster analysis on data in order to segment data into appropriate clusters for subsequent processing (see [0010], lines 5-8), including the further limitations of determining from the statistical data whether each of the one or more clusters is abnormal when compared to a predefined value (see [0217]); and reporting [classifying] at least one of the one or more clusters as an abnormal cluster of objects in the data stream (see [0217]) in order to improve the capability of an intrusion detection algorithm to be scalable and efficient in the handling data in real-time systems.

It would have been obvious to one of ordinary skill in the to use the features of determining whether an entire cluster is abnormal and reporting that abnormality as disclosed by Lundahl using the statistical data determined by Ye. One would have been motivated to do so in order to improve the capability of an intrusion detection algorithm to be scalable and efficient in the handling data in real-time systems (Ye: see [0010], lines 6-8).

Referring to claim 2, the combination of Ye and Lundahl (hereafter Ye/Lundahl) discloses the method of claim 1, wherein the step of creating one or more clusters further comprises:

computing one or more similarity values for a given object relating to one or more existing clusters (Ye: see [0157]-[0162]); and

determining a closest cluster for the object based on the one or more similarity values (Ye: see [0163]).

Referring to claim 3, Ye/Lundahl discloses the method of claim 2, further comprising the steps of:

determining whether to add the object to the closest cluster (Ye: see [0157-[0163]);

adding the object to the closest cluster when determined and updating the statistical data of the closest cluster (Ye: see [0041]-[0042]); and

creating a new cluster comprising the object when the object is not added to the closest cluster (Smolsky: see column 13, lines 31-35), and generating statistical data of the new cluster (Smolsky: see column 9, lines 4-14 and column 14, lines 53 – column 15, line 2).

Referring to claim 4, Ye/Lundahl discloses the method of claim 3, wherein the step of determining whether to add the object to the closest cluster further comprises the step of determining if the similarity value is greater than a user-defined threshold (Ye: see [0173]).

Referring to claim 5, Ye/Lundahl discloses the method of claim 1, wherein the step of determining from the statistical data whether each of the one or more clusters is abnormal further comprises the steps of:

determining which clusters present at a first time were not present at a second time, wherein the second time is before the first time; determining which of the clusters, present at the first time and not present at the second time, contain fewer than a user-

defined number of objects; and reporting clusters with fewer than the user-defined number of objects as abnormalities (Lundahl: see [0217]).

Referring to claim 6, Ye/Lundahl discloses the method of claim 1, wherein the statistical data of each cluster is stored using an incremental updating process (Ye: see [0154], lines 8-15).

Referring to claim 9, Ye/Lundahl discloses the method of claim 1, wherein the statistical data of each cluster comprises a number of objects [number of data points] in each cluster (Ye: see [0154], lines 9-13).

Referring to claim 11, Ye/Lundahl discloses the method of claim 1, wherein the step of creating one or more clusters further comprises the step of applying one or more weights to one or more attributes (Ye: see [0174]).

Referring to claim 12, Ye/Lundahl discloses the method of claim 1, wherein abnormalities comprise intrusions in a network (Ye: see [0030], lines 10-17).

Referring to claim 16, Ye discloses an apparatus for monitoring abnormalities in a data stream (see abstract and [0030]), comprising:

a memory (digital storage medium) (see [0026] and Fig 1); and

at least one processor [computer system] coupled [network] to a memory and operative to:

(i) receive a plurality of objects in the data stream [stream of data] (see [0035], lines 5-8) and

(ii) create one or more clusters from the plurality of objects (see [0035], lines 10-13), wherein at least a portion of the one or more clusters comprise

statistical data [sample variance, sample covariance and sample mean] of the respective cluster (see [0041]).

Ye discloses clustering objects and determining if an object is abnormal compared to a distance value (see [0157]-[0170]), however, Ye fails to explicitly disclose the further limitation of (iii) determine from the statistical data whether each of the one or more clusters is abnormal when compared to a predefined value. Lundahl discloses performing cluster analysis on data in order to segment data into appropriate clusters for subsequent processing (see [0010], lines 5-8), including the further limitation of (iii) determine from the statistical data whether each of the one or more clusters is abnormal when compared to a predefined value (see [0217]) in order to improve the capability of an intrusion detection algorithm to be scalable and efficient in the handling data in real-time systems.

It would have been obvious to one of ordinary skill in the to use the feature of determining whether an entire cluster is abnormal and as disclosed by Lundahl using the statistical data determined by Ye. One would have been motivated to do so in order to improve the capability of an intrusion detection algorithm to be scalable and efficient in the handling data in real-time systems (Ye: see [0010], lines 6-8).

Referring to claim 17, Ye/Lundahl discloses the apparatus of claim 16, wherein the operation of creating one or more clusters further comprises:

computing one or more similarity values for a given object relating to one or more existing clusters (Ye: see [0157]-[0162]); and

determining a closest cluster for the object based on the one or more similarity values (Ye: see [0163]).

Referring to claim 18, Ye/Lundahl discloses the apparatus of claim 17, further comprising:

determining whether to add the object to the closest cluster (Ye: see [0157-[0163]);

adding the object to the closest cluster when determined and updating the statistical data of the closest cluster (Ye: see [0041]-[0042]); and

creating a new cluster comprising the object when the object is not added to the closest cluster (Smolsky: see column 13, lines 31-35), and generating statistical data of the new cluster (Smolsky: see column 9, lines 4-14 and column 14, lines 53 – column 15, line 2).

Referring to claim 19, Ye/Lundahl discloses the apparatus of claim 18, wherein determining whether to add the object to the closest cluster further comprises the step of determining if the similarity value is greater than a user-defined threshold (Ye: see [0173]).

Referring to claim 20, Ye/Lundahl discloses the apparatus of claim 17, wherein the operation of determining from the statistical data whether each of the one or more clusters is abnormal further comprises:

determining which clusters present at a first time were not present at a second time, wherein the second time is before the first time; determining which of the clusters, present at the first time and not present at the second time, contain fewer than a user-

defined number of objects; and reporting clusters with fewer than the user-defined number of objects as abnormalities (Lundahl: see [0217]).

Referring to claim 21, Ye/Lundahl discloses the apparatus of claim 16, wherein the statistical data of each cluster is stored using an incremental updating process (Ye: see [0154], lines 8-15).

Referring to claim 24, Ye/Lundahl discloses the apparatus of claim 16, wherein the statistical data of each cluster comprises a number of objects [number of data points] in each cluster (Ye: see [0154], lines 9-13).

Referring to claim 26, Ye/Lundahl discloses the apparatus of claim 16, wherein the operation of creating one or more clusters further comprises the step of applying one or more weights to one or more attributes (Ye: see [0174]).

Referring to claim 27, Ye/Lundahl discloses the apparatus of claim 16, wherein abnormalities comprise intrusions in a network (Ye: see [0030], lines 10-17).

Referring to claim 31, Ye discloses an article of manufacture for monitoring abnormalities in a data stream (see abstract and [0030]), comprising a machine readable containing one or more programs which when executed implement the steps of:

receiving a plurality of objects in the data stream [stream of data] (see [0035], lines 5-8); and

creating one or more clusters from the plurality of objects (see [0035], lines 10-13), wherein at least a portion of each of the one or more clusters comprises statistical

data [sample variance, sample covariance and sample mean] representative of the respective cluster (see [0041]).

Ye discloses clustering objects and determining if an object is abnormal compared to a distance value (see [0157]-[0170]), however, Ye fails to explicitly disclose the further limitation of determining from the statistical data whether each of the one or more clusters is abnormal when compared to a predefined value.

Lundahl discloses performing cluster analysis on data in order to segment data into appropriate clusters for subsequent processing (see [0010], lines 5-8), including the further limitation of determining from the statistical data whether each of the one or more clusters is abnormal when compared to a predefined value (see [0217]) (see [0217]) in order to improve the capability of an intrusion detection algorithm to be scalable and efficient in the handling data in real-time systems.

It would have been obvious to one of ordinary skill in the to use the feature of determining whether an entire cluster is abnormal as disclosed by Lundahl using the statistical data determined by Ye. One would have been motivated to do so in order to improve the capability of an intrusion detection algorithm to be scalable and efficient in the handling data in real-time systems (Ye: see [0010], lines 6-8).

10. Claims 7, 10, 22 and 25 are rejected under 35 U.S.C. 103(a) as being unpatentable over US PGPub 2002/0161763 to Ye et al in view of US PGPub 2002/0107858 to Lundahl et al as applied respectively to claims 1 and 16 above, and further in view of US Patent No 6,625,585 to MacCuish et al (hereafter MacCuish et al).

Referring to claim 7, Ye/Lundahl discloses statistical data. However, Ye/Lundahl fails to explicitly disclose the further limitation wherein the statistical data of each cluster comprises one or more statistical counts of each pairwise attribute. MacCuish et al disclose clustering data (see abstract) including the further limitation wherein the statistical data of each cluster comprises one or more statistical counts of each pairwise attribute (see column 14, lines 44-62) so in order to improve the accuracy of calculating the similarity of the clusters.

It would have been obvious to one of ordinary skill in the art at the time the invention was made to utilize pairwise attributes of MacCuish et al as the type of statistical data utilized by Ye/Lundahl. One would have been motivated to do so in order to improve the accuracy of calculating the similarity of the clusters.

Referring to claim 10, Ye/Lundahl discloses statistical data. However, Ye/Lundahl fails to explicitly disclose the further limitation wherein the statistical data is stored periodically at intervals chosen based on a pyramidal distribution. MacCuish et al disclose clustering data (see abstract) including the further limitation wherein the statistical data is stored periodically at intervals chosen based on a pyramidal distribution (see column 14, lines 27-29) since the data being clustered is being

transmitted in a stream which means that new data is constantly being clustered and clustering at a periodic interval decreases utilized system resources.

It would have been obvious to one of ordinary skill in the art at the time the invention was made to utilize the feature of periodically storing the statistics of MacCuish et al as the type of statistical data utilized by Ye/Lundahl. One would have been motivated to do so since the data being clustered is being transmitted in a stream, which means that new data is constantly being clustered and clustering at a periodic interval decreases utilized system resources.

Referring to claim 22, the claim is rejected on the same grounds as claim 7.

Referring to claim 25, the claim is rejected on the same grounds as claim 10.

11. Claims 8, 13-15, 23 and 28-30 are rejected under 35 U.S.C. 103(a) as being unpatentable over US PGPub 2002/0161763 to Ye et al in view of US PGPub 2002/0107858 to Lundahl et al as applied respectively to claims 1 and 16 above, and further in view of US PGPub 2004/0098617 to Sekar (hereafter Sekar).

Referring to claim 8, Ye/Lundahl discloses statistical data of each cluster. However, Ye/Lundahl fails to explicitly disclose the further limitation wherein the statistical data of each cluster comprises one or more statistical counts of each categorical attribute. Sekar discloses statistical data, including the further limitation wherein the statistical data of each cluster comprises one or more statistical counts of each categorical attribute (Sekar: see [0088], lines 1-7) in order to increase the speed and efficiency at which intrusions can be detected in a large sample of data.

It would have been obvious to one of ordinary skill in the art to use the statistical counts of Sekar as additional data to the statistical data Ye/Lundahl. One would have been motivated to do so in order to increase the speed and efficiency at which intrusions can be detected in a large sample of data.

Referring to claim 13, Ye/Lundahl discloses abnormalities, which represent intrusions in a network. However, Ye/Lundahl fail to explicitly disclose the further limitation of wherein the step of receiving a plurality of objects further comprises the step of collecting source IP (Internet Protocol) address data, destination IP address data and signature data. Sekar discloses determining abnormalities in data, wherein abnormalities comprise intrusions in a network (see abstract), including the further limitation of a step of receiving a plurality of objects which comprises a step of collecting source IP (Internet Protocol) address data [source address], destination IP address data [destination address] and signature data (Smolsky: see column 5, line 34 and line 45) in order to increase the speed and efficiency at which intrusions can be detected in a large sample of data.

It would have been obvious to one of ordinary skill in the art to use the IP address and signature data collected by Sekar with the data of Ye/Lundahl in order to determine the intrusions in a network. One would have been motivated to do so in order to increase the speed and efficiency at which intrusions can be detected in a large sample of data.

Referring to claim 14, Ye/Lundahl discloses abnormalities, which represent intrusions in a network and the step of clustering data. However, Ye/Lundahl fail to

explicitly disclose the further limitation of wherein the step of creating one or more clusters further comprises the step of clustering source IP address data, destination IP address data and signature data. Sekar discloses determining abnormalities in data, wherein abnormalities comprise intrusions in a network (see abstract), including collecting source IP (Internet Protocol) address data [source address], destination IP address data [destination address] and signature data (Smolsky: see column 5, line 34 and line 45) in order to increase the speed and efficiency at which intrusions can be detected in a large sample of data.

It would have been obvious to one of ordinary skill in the art to use the IP address and signature data collected by Sekar as the data being clustered by Ye/Lundahl. One would have been motivated to do so in order to increase the speed and efficiency at which intrusions can be detected in a large sample of data.

Referring to claim 15, Ye/Lundahl discloses abnormalities, which represent intrusions in a network and the step of determining from statistical data whether abnormalities exist. However, Ye/Lundahl fail to explicitly disclose the further limitation of wherein the step of determining from the statistical data whether each of the one or more clusters is abnormal comprises the step of detecting one or more intrusions from statistical data of source IP address data, destination IP address data and signature data. Sekar discloses determining abnormalities in data, wherein abnormalities comprise intrusions in a network (see abstract), including wherein the step of determining from the statistical data whether one or more abnormalities exist further comprises the step of detecting one or more intrusions from statistical data of source IP

address data, destination IP address data and signature data (Sekar: see [0032]) in order to increase the speed and efficiency at which intrusions can be detected in a large sample of data.

It would have been obvious to one of ordinary skill in the art to use the IP address and signature data collected by Sekar as the data being clustered by Ye/Lundahl. One would have been motivated to do so in order to increase the speed and efficiency at which intrusions can be detected in a large sample of data.

Referring to claim 23, the claim is rejected on the same grounds as claim 8.

Referring to claim 28, the claim is rejected on the same grounds as claim 13.

Referring to claim 29, the claim is rejected on the same grounds as claim 14.

Referring to claim 30, the claim is rejected on the same grounds as claim 15.

Response to Arguments

12. Applicant's arguments with respect to claims 1-31 have been considered but are moot in view of the new ground(s) of rejection.

Conclusion

13. The prior art made of record and not relied upon is considered pertinent to applicant's disclosure.

- The article "COOLCAT: An entropy-based algorithm for categorical clustering" to Barbara et al.

Contact Information

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Kimberly Lovel whose telephone number is (571) 272-2750. The examiner can normally be reached on 8:00 - 4:00.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, John Cottingham can be reached on (571) 272-7079. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

Kimberly Lovel
Examiner
Art Unit 2167

26 February 2007
kml

JOHN COTTINGHAM
SUPERVISORY PATENT EXAMINER
TECHNOLOGY CENTER 2100